Project 1: Test a Perceptual Phenomenon

Feb 2, 2016 - resubmission

By Teresa Aysan

<u>1. What is our independent variable? What is our dependent variable?</u>

The independent variable is the two conditions: a congruent words condition, and an incongruent words condition.

The dependent variable is the time it takes to name the ink colors.

<u>2. What is an appropriate set of hypotheses for this task? What kind of statistical test do you expect to perform? Justify your choices.</u>

<u>2a) Appropriate set of hypotheses:</u> Null and alternative hypotheses are clearly stated in words and mathematically. Symbols in the mathematical statement are defined.

- Test the research hypothesis that it will take more than 1.6 times longer for the participants to say the list of incongruent words than to say the list of congruent words, with a confidence level of 0.05.

  i = time to say list of incongruent words

  c = time to say list of congruent words

  H_o = Null Hypothesis

  H_1 = Alternate Hypothesis

  μ = population mean (expected value) of two populations i and c.

  α = Alpha – the confidence level: *also called the "significance level." By* definition*, the* alpha *level is the probability of rejecting the null hypothesis when the null hypothesis is true.*[22]

- The Null Hypothesis is that it will take 1.6 times longer for the participants to say the list of incongruent words than to say the list of congruent words. H_o will be measured in Difference (D) between the two samples.

  $\boxed{H_o\text{: } \mu_0 = (\mu_c * 1.6) - \mu_c}$

  The Alternative Hypothesis is that it will take more than 1.6 times longer for the participants to say the list of incongruent words than to say the list of congruent words, with a confidence level of 0.05. The Alternative Hypothesis will also be measured in Difference.

H_1: $\mu_i > (\mu_c*1.6)-\mu_c$

$\alpha = 0.05$

- *Based on the given random sample from the population: if the sample data are consistent with the null hypothesis, then [do not reject]() the null hypothesis; if the sample data are inconsistent with the null hypothesis, then reject the null hypothesis and conclude that the alternative hypothesis is true.*[15]

2b) Justify choice of hypotheses:

- I chose to do an upper-tailed test (>) for the following reasons:
  - I did not want to do a $\mu_i = \mu_c$ test because I know from taking the test that is very unlikely that it takes the same amount of time to say the incongruent words as the congruent words.
  - I did not want to do a $\mu_i < \mu_c$ test because common sense dictates that it is not probable that it will take less time to say the incongruent words than the congruent words.
  - I did not want to do a simple a $\mu_i > \mu_c$ because common sense dictates that it is obviously true that it takes longer to say the incongruent words than to say the congruent words.
  - So, I know that μ_i will take longer, but how much longer? I estimated 1.6 times longer based on how much longer it took me to say the words.

2c) Type of Statistical test that I will perform: Spec: A statistical test has been proposed which will distinguish the proposed hypotheses. Any assumptions made by the statistical test are addressed.

- I will perform an upper-tailed **dependent** t-test for paired samples[23], with alpha value of 0.05.

  Paired samples *t*-tests typically consist of a sample of matched pairs of similar units, or one group of units that has been tested twice (a "repeated measures" *t*-test). … Paired samples *t*-tests are often referred to as "dependent samples *t*-tests".[23]

  In the case of the Stroop Effect data, the same people were tested twice with two different sets of data (congruent and incongruent), and the amount of time taken for each test was recorded for each person. The congruent and incongruent times were paired for each person.

- $$t = \frac{\overline{X}_D - \mu_0}{\frac{s_D}{\sqrt{n}}}.$$

For this equation, the differences between the time taken for the incongruent words test and the congruent words test for all pairs must be calculated.

The average (X_D) and standard deviation (s_D) of those differences are used in the equation. The constant μ_0 is non-zero (in this case based on μ_i being 1.6x longer than μ_c) since I want to test whether the average of the difference is significantly different from μ_0. The degree of freedom used is $n - 1$, where n represents the number of pairs.[23]

- I chose the t-test because the sample size is 24 which is considered small (<30).[18]

<u>3. Report some descriptive statistics regarding this dataset. Include at least one measure of central tendency and at least one measure of variability.</u>

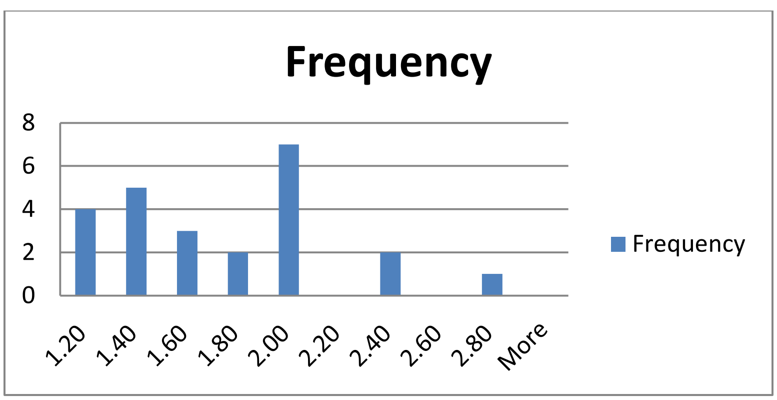
<u>3a) Central Tendency</u>: "The term 'measures of central tendency' refers to finding the mean, median and mode." [5]

- Mean(D) = 7.96
- Where D = difference between the two sample data

<u>3b) Variability</u>: "Statisticians use summary measures to describe the amount of variability or spread in a set of data. The most common measures of variability are the range, the interquartile range (IQR), variance, and standard deviation." [3]

- Variance(D) = 23.67
- Standard Deviation(D) = 4.86

<u>4. Provide one or two visualizations that show the distribution of the sample data. Write one or two sentences noting what you observe about the plot or plots.</u>



The histogram shows that the mode is 2 (2x longer to say the incongruent words than the congruent words): When the bins are defined as above, the most common time is 2x longer.

The histogram shows that most people take ≤ 2x longer to say the incongruent words than the congruent words.

5. Now, perform the statistical test and report your results. What is your confidence level and your critical statistic value? Do you reject the null hypothesis or fail to reject it? Come to a conclusion in terms of the experiment task. Did the results match up with your expectations?

I am using Dependent *t*-test for paired samples[23]

$$t = \frac{\overline{X}_D - \mu_0}{\frac{s_D}{\sqrt{n}}}$$

Dataset – time it takes to say a list of congruent and incongruent words:

dataset:	Congruent time	Incongruent time	Difference	Deviation ^2	x longer
	12.079	19.278	7.199	0.586	1.596
	16.791	18.741	1.950	36.178	1.116
	9.564	21.214	11.650	13.581	2.218
	8.630	15.687	7.057	0.824	1.818
	14.669	22.803	8.134	0.029	1.555
	12.238	20.878	8.640	0.456	1.706
	14.692	24.572	9.880	3.668	1.672
	8.987	17.394	8.407	0.196	1.935
	9.401	20.762	11.361	11.534	2.208
	14.480	26.282	11.802	14.724	1.815
	22.328	24.524	2.196	33.279	1.098
	15.298	18.644	3.346	21.333	1.219
	15.073	17.510	2.437	30.556	1.162
		20.330			

	16.929		3.401	20.828	1.201
	18.200	35.255	17.055	82.632	1.937
	12.130	22.158	10.028	4.257	1.827
	18.495	25.139	6.644	1.744	1.359
	10.639	20.429	9.790	3.331	1.920
	11.344	17.425	6.081	3.549	1.536
	12.369	34.288	21.919	194.720	2.772
	12.944	23.894	10.950	8.911	1.846
	14.233	17.960	3.727	17.959	1.262
	19.710	22.058	2.348	31.548	1.119
	16.004	21.157	5.153	7.906	1.322
Sum	337.23	528.38	191.16	544.33	
Mean (Avg)	14.05	22.02	7.96	22.68	
N = count			24		
D = Difference					
n = N-1 (degree of freedom)				23	
Mean(D)				7.96	
Sample Variance(D) = average squared deviation from the mean - using n				23.67	
S(D): Sample Std Dev = SqRt(Var)				4.86	
one sd above mean				12.83	
one sd below mean				3.10	
Standard Error (SE) = Std Dev / SQRT(n)				1.01	
Factor chosen for my test				1.6	times longer
μc = mean of congruent time					

μi = mean of incongruent time					
$\mu 0 = (\mu c*1.6)-\mu c$				8.43	
I want to test whether the average of the difference is significantly different from μ_0.					
t statistic:				(0.46)	
$$t = \frac{\overline{X}_D - \mu_0}{\frac{s_D}{\sqrt{n}}}.$$				0.46	absolut e value
		The t statistic is the number of std deviations			
		the sample average is			
		from the hypothesized mean			
alpha - one tail significance level (chosen by me)				0.05	
Confidence level				95%	
For an upper, one-sided test, find the t table column corresponding to 1-α and					
reject the null hypothesis if the test statistic is greater than the table value. [30]					
df				23	
alpha - one tail significance level (chosen by me)				0.05	
Critical statistic value from t table				1.714	
The t statistic is less than the critical statistic value,					
therefore failed to reject the null hypothesis.					

The final conclusion is to fail to reject the null hypothesis because the sample data are not very unlikely.

We have not shown that it will take more than 1.6 times longer for the participants to say the list of incongruent words than to say the list of congruent words, with a confidence level of 0.05.

<u>Did the results match up with your expectations?</u>

The spread of the frequency in the histogram was close to what I expected.

However, I expected that we would reject the null hypothesis. I learned that an alpha level of 0.05 allows for quite a large spread before the null hypothesis can be rejected.

6. Optional: What do you think is responsible for the effects observed? Can you think of an alternative or similar task that would result in a similar effect? Some research about the problem will be helpful for thinking about these two questions!

I chose not to do the optional question.

Resources:

1. Link from the project overview: Stroop Effect
2. *Link to page with symbols: http://brownmath.com/swt/symbol.htm*
3. Measure of variability: https://www.google.ca/?gws_rd=cr&ei=oBduVsjmBobGmQHttLHQCw#q=measure+of+variability+in+statistics
4. Measure of central tendency: https://www.google.ca/?gws_rd=cr&ei=oBduVsjmBobGmQHttLHQCw#q=measure+of+central+tendency
5. RegentsPrep Measure of central tendency: http://www.regentsprep.org/regents/math/algebra/ad2/measure.htm
6. How do you calculate variance in Excel?: http://www.investopedia.com/ask/answers/041615/how-do-you-calculate-variance-excel.asp
7. Calculating Variance and Standard Deviation in 4 Easy Steps: http://www.macroption.com/calculate-variance-standard-deviation-4-steps/
8. Population vs. Sample Variance and Standard Deviation http://www.macroption.com/population-sample-variance-standard-deviation/
9. How to Calculate Mean, Standard Deviation, and Standard Error: http://www.wikihow.com/Calculate-Mean,-Standard-Deviation,-and-Standard-Error
10. How to Calculate a Standard Error of the Mean in Excel: https://www.google.ca/?gws_rd=cr,ssl&ei=z_ueVtzrHqmHjgTNsp-YCQ#q=how+to+calculate+the+standard+error+in+excel
11. Standard error https://en.wikipedia.org/wiki/Standard_error
12. Z-Score: Definition, Formula and Calculation: http://www.statisticshowto.com/how-to-calculate-a-z-score/
13. how to read the z score table:

http://www.dummies.com/how-to/content/how-to-find-probabilities-for-z-with-the-ztable.html

14. Hypothesis Testing: Upper-, Lower, and Two Tailed Tests: http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_HypothesisTest-Means-Proportions/BS704_HypothesisTest-Means-Proportions3.html

15. What is a null hypothesis:
    https://en.wikipedia.org/wiki/Null_hypothesis
    http://study.com/academy/lesson/what-is-a-null-hypothesis-definition-examples.html
    Usually the null hypothesis is a statement of 'no effect' or 'no difference'. It is often symbolized as H_0.
    The statement that is hoped or expected to be true instead of the null hypothesis is the **alternative hypothesis**.[4] Symbols include H_1 and H_a.
    Statistical significance test: "Very roughly, the procedure for deciding goes like this: Take a random sample from the population. If the sample data are consistent with the null hypothesis, then do not reject the null hypothesis; if the sample data are inconsistent with the null hypothesis, then reject the null hypothesis and conclude that the alternative hypothesis is true."
    According to this view, the null hypothesis must be numerically exact—it must state that a particular quantity or difference is equal to a particular number. In classical science, it is most typically the statement that there is *no effect* of a particular treatment; in observations, it is typically that there is *no difference* between the value of a particular measured variable and that of a prediction.
    To overcome any possible ambiguity in reporting the result of the test of a null hypothesis, it is best to indicate whether the test was two-sided and, if one-sided, to include the direction of the effect being tested.

16. What are the kinds of statistical tests:
    https://cyfernetsearch.org/types-statistical-tests

17. Z-test
    https://en.wikipedia.org/wiki/Z-test

18. t-test:
    https://www.google.ca/?gws_rd=cr&ei=gU-gVorjG8b5jgSeppCIDA#q=what+is+a+t+test
19. t-test:
    http://blog.minitab.com/blog/statistics-and-quality-data-analysis/what-is-a-t-test-and-why-is-it-like-telling-a-kid-to-clean-up-that-mess-in-the-kitchen
20. t-test:
    http://blog.minitab.com/blog/statistics-and-quality-data-analysis/what-are-t-values-and-p-values-in-statistics

21. How to create a histogram in Excel
    http://www.statisticshowto.com/frequency-distribution-table-in-excel/

22. Definition of Alpha: https://www.google.com/?gws_rd=ssl#q=what+is+alpha+in+statistics+and+what+does+it+mean

23. Unpaired and paired 2 sample t-tests: https://en.wikipedia.org/wiki/Student%27s_t-test#Unpaired_and_paired_two-sample_t-tests

24. Means – expected value https://en.wikipedia.org/wiki/Expected_value

25. Calculating P value
http://trendingsideways.com/index.php/the-p-value-formula-testing-your-hypothesis/

If we get a P-value smaller than our significance level, we can reject the null hypothesis.
The t statistic is the number of standard deviations your sample average is from the hypothesized mean.
In the case of the t statistic, the number of degrees of freedom is just one less than the sample size: **n-1**. (This is because we have used our sample data twice, once to find the sample mean, and again to find the sample standard deviation.)
Once we know our t statistic and our degrees of freedom, we can just plug them into a table or statistical software.

26. T table:
https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg

27. Reading t table
https://ca.answers.yahoo.com/question/index?qid=20080814081530AAnwu75

The top numbers 0.05 , 0.02, 0.01 are significance levels. The problem must specify what the significance level is (usually, not always, it is 0.05 or 0.01). Read the t-value under that column.
If you don't find a t-value, take the closest t-value.
For example, in the attached table, if the degree of freedom is 63, you can use t=60 or t=65.

If your t value is 2.3197 , you should first know your level of significance (you or the questioner decides this).
If it is a one-tailed test, look under 0.05 directly (the attached table tells you whether it is one or two-tailed but most tables won't).
The critical t value is 1.653. Your computed t exceeds the critical t, so you reject your null hypothesis (yes, even if it exceeds by a small amount, you still reject it).

If it is two-tailed and the level of significance is 0.05, look under 0.025 (one half of 0,.05).

How do you know whether it is one-tailed or two-tailed. This must be decided by reading the question.

H0: $\mu(1) = \mu(2)$
HA:: $\mu(1) > \mu(2)$ (one-tailed)
HA:: $\mu(1) < \mu(2)$ (one-tailed)
HA:: $\mu(1) \neq \mu(2)$ (two-tailed)

28. Calculating Confidence level

http://www.itl.nist.gov/div898/handbook/prc/section1/prc14.htm
Confidence intervals are constructed at a *confidence level*, such as 95 %, selected by the user. What does this mean? It means that if the same population is sampled on numerous occasions and interval estimates are made on each occasion, the resulting intervals would bracket the true population parameter in approximately 95 % of the cases. A confidence stated at a $1-\alpha$ level can be thought of as the inverse of a significance level, α.

29. Finding Critical Statistic value – Z table
    http://www.statisticshowto.com/find-a-critical-value/
    http://www.math.armstrong.edu/statsonline/5/5.3.2.html

30. Critical value – T table
    http://www.itl.nist.gov/div898/handbook/eda/section3/eda3672.htm
    If the absolute value of the test statistic is greater than the critical value…
    For an upper, one-sided test, find the column corresponding to 1-α and reject the null hypothesis if the test statistic is greater than the table critical value.